

fragments or words to form a proposed word included in a backup dictionary of the speech recognition system using a spelling rule associated with the word fragment. As a result of using the associated spelling rule, a spelling of the proposed word differs from a spelling that would result from merely concatenating the particular word fragment with the one or more adjacent word fragments or words.

If the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, the recognition candidate is modified to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word. If the word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, the recognition candidate is discarded.

Independent claim 46 relates to a computer-implemented speech recognition system that uses an expanded effective active vocabulary. The system includes a storage device configured to store an active vocabulary that includes multiple entries corresponding to words, commands, and word fragments. The system includes a processor configured to substantially perform the method of claim 1.

Independent claim 51 relates to computer software, residing on a computer readable medium, for a speech recognition system that uses an expanded effective active vocabulary to recognize words and commands. The computer software includes instructions for causing a computer to substantially perform the method of claim 1.

Applicant requests withdrawal of the rejection of claims 1, 46, and 51 because neither Sherwood, Roberts, Baker, nor any proper combination of the three describes or suggests (1) combining a particular word fragment with one or more adjacent word fragments using an associated spelling rule to form a proposed word having a spelling that differs from a spelling that would result from merely concatenating the word fragment with the one or more adjacent word fragments or words or (2) when the word fragment may not be combined with one or more

adjacent word fragments or words to form a proposed word, discarding the recognition candidate that includes the word fragment.

Recognizing that Sherwood and Roberts are silent as to the use of spelling rules, the Examiner relies on Baker:

In a similar field of endeavor, Baker teaches a method of adding a word to a speech recognition vocabulary by creating a collection of possible phonetic pronunciation from a spelling of a word, wherein the collection is created by comparing the spelling to a rules list of letter strings associated with phones at col. 15, line 56 continuing to col. 18, line 26.

Therefore, it would have been obvious to one of ordinary skill at the time of the invention to modify the recognition system of Sherwood to implement spelling rules for adding new words to the system vocabulary as taught by Baker, for the purpose of ensuring that new words to be added to the vocabulary are actually valid words.

Even assuming for sake of argument that the Examiner's conclusions are correct, this in no way establishes that Sherwood, Roberts and Baker could have been combined to produce the subject matter of claims 1, 46 and 51. In particular, those claims do not recite either the mere use of spelling rules or the use of spelling rules to add new words to a system vocabulary, as the Examiner argues is taught by Baker. Rather, they specifically recite the use of a spelling rule to produce a proposed word having a spelling that differs from a spelling that would result from merely concatenating a particular word fragment with the one or more adjacent word fragments or words, which is not described or suggested by Baker (or by Sherwood or Roberts).

Baker is entirely silent as to the concept of producing a word having a spelling that differs from the spelling that would result from concatenating a word fragment with one or more adjacent word fragments or words. Indeed, Baker's system uses the spelling of the word as a starting point. See Baker at col. 15, lines 57-60 ("the user types in the word (step 1305) ...") and col. 17, lines 32-33 ("after the user spells the word (step 1505) ..."). Accordingly, Baker cannot remedy the failure of Sherwood and Roberts to describe or suggest combining a particular word fragment with one or more adjacent word fragments using an associated spelling rule to form a proposed word having a spelling that differs from a spelling that would result from merely concatenating the word fragment with the one or more adjacent word fragments or words, as recited in each of independent claims 1, 46 and 51.

Sherwood, Roberts and Baker also are silent with respect to discarding a recognition candidate if a word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word. As best understood, the Examiner appears to assert that Roberts describes this aspect of the claims. However, Roberts, which is directed to adding words to a recognition system vocabulary, is silent as to when recognition candidates that include word fragments are used or discarded. Sherwood and Baker similarly fail to describe or suggest this aspect of the claims.

For at least the reasons provided above, claims 1, 46, and 51 are allowable over Sherwood in view of Roberts and Baker, as are their dependent claims.

Claim 37

Independent claim 37 relates to a method of generating acoustic models of word fragments. The method includes comparing words of an active vocabulary to similar words of a backup dictionary to identify spelling rules that may be used to convert the words of the active vocabulary to words of the backup dictionary, and employing the spelling rules in identifying word fragments.

Applicant requests withdrawal of the rejection of claim 37 because any possible combination of Sherwood, Roberts and Baker would still fail to describe or suggest comparing words of an active vocabulary with similar words of a backup dictionary to identify spelling rules that may be used to convert the words of the active vocabulary to words of the backup dictionary and employing the spelling rules in identifying word fragments.

As previously discussed, Sherwood and Roberts are silent as to identifying spelling rules and using spelling rules to identify word fragments. Furthermore, even assuming for sake of argument that Baker describes the use of spelling rules, Baker nowhere describes or suggests identifying spelling rules that may be used to convert words of an active vocabulary to words of a backup dictionary, and Baker also nowhere describes or suggests employing the spelling rules to identify word fragments, as recited in claim 37. As discussed above, in Baker's system, a user spells and speaks a word to be added to a vocabulary and the spoken word is recognized against

the constraint grammar. Baker's speech recognition system does not identify spelling rules to convert words from an active vocabulary to words of a backup dictionary, and does not employ such spelling rules to identify fragments of words. Accordingly, claim 37 is allowable over Sherwood in view of Roberts and Baker.

Claim 34

Independent claim 34 relates to a method of recognizing speech. Recognition candidates are received from a speech recognizer that uses a set of acoustic models representative of an active vocabulary. The set of acoustic models includes models of words, models of roots that are not words, and models of affixes that are not words. The affixes include prefixes and suffixes.

When a received recognition candidate includes an affix, the method includes combining the affix with one or more adjacent words, roots, or other affixes to form a new word and modifying the recognition candidate to substitute the new word for the affix and the one or more adjacent words, roots, or other affixes used to form the new word.

Formation of the new word includes using a spelling rule associated with the affix that causes the spelling of the new word to differ from a spelling that would result from merely concatenating the affix with the one or more adjacent words, roots, or other affixes.

As discussed above with respect to claim 1, neither Sherwood, Roberts, Baker, nor any combination of the three describes or suggests combining a particular word fragment with one or more adjacent word fragments using an associated spelling rule in forming the proposed word and a spelling of a proposed word differing from a spelling that would result from merely concatenating the word fragment with the one or more adjacent word fragments or words. For this reason, the combination of Sherwood, Roberts and Baker also fails to describe or suggest formation of a new word using a spelling rule associated with an affix that causes the spelling of the new word to differ from a spelling that would result from merely concatenating the affix with the one or more adjacent words, root, or other affixes that that form the new word.

Kanevsky does not remedy the failure of Sherwood, Roberts and Baker to describe or suggest combining a particular word fragment with one or more adjacent word fragments using

an associated spelling rule in forming the proposed word and a spelling of a proposed word differing from a spelling that would result from merely concatenating the word fragment with the one or more adjacent word fragments or words.

Kanevsky relates to a language model that is constructed by dividing words into word components. See Kanevsky at col. 2, lines 26-33. Kanevsky generates the language model by calculating the probability of a word as a "weighted sum of output probabilities of several" language models, which are built on word components and combinations of word components. See Kanevsky at col. 3, lines 6-17.

Though Kanevsky discusses word components, Kanevsky fails to show combining word components to form a proposed word using a spelling rule that produces a spelling of the proposed word that differs from a spelling that would result from merely concatenating the word component with other word components. Rather, as Kanevsky explains, the word components are merely concatenated: "stems are connected consequently ... and matched ... to see which concatenations of stems produce existing words in the vocabulary." See Kanevsky at col. 7, lines 16-21 and Fig. 7. For at least this reason, claim 34 is allowable over Sherwood in view of Roberts, Baker and Kanevsky.

Claims 3-9, 22-29 and 38-45

Applicant requests reconsideration and withdrawal of the rejection of claims 3-9, 22-29 and 38-45, which depend from claims 1 and 37. In particular, for the reasons discussed above with respect to claim 34, Kanevsky does not remedy the failure of Sherwood, Roberts and Baker to describe or suggest the subject matter of independent claims 1 and 37.

Claim 35

Independent claim 35 relates to a method of generating an acoustic model of a word fragment. The method includes comparing a word of an active vocabulary to a similar word of a backup dictionary to identify a word fragment that may be used to convert the word of the active vocabulary to the word of the backup dictionary. The method also includes generating the

acoustic model of the word fragment using a portion of an acoustic model of the word of the backup dictionary that is not included in an acoustic model of the word of the active vocabulary.

Applicant requests withdrawal of the rejection of claim 35 as being obvious over Sherwood in view of Roberts because any theoretical combination of Sherwood and Roberts would still fail to describe or suggest comparing a word of an active vocabulary to a similar word of a backup dictionary to identify a word fragment that may be used to convert the word of the active vocabulary to the word of the backup dictionary, and would also fail to describe or suggest generating the acoustic model of the word fragment using a portion of an acoustic model of the word of the backup dictionary that is not included in an acoustic model of the word of the active vocabulary, as recited in claim 35.

As the Examiner correctly points out, Sherwood simply does not relate to word fragments. Accordingly, Sherwood cannot describe or suggest identifying a word fragment and generating an acoustic model of the word fragment.

Moreover, Roberts fails to cure the deficiencies of Sherwood. Roberts' speech recognition system does not identify a word fragment that may be used to convert a word of an active vocabulary to a word of a backup dictionary, and does not generate an acoustic model of the word fragment using a portion of an acoustic model of the word of the backup dictionary that is not included in the acoustic model of the word of the active vocabulary. While, as the Examiner notes, Roberts teaches the identification of fragments or segments of a phrase "This is a black cow," Roberts does so by analyzing the speech frames of an utterance, and does not do so by comparing the speech frames of a word from an active vocabulary to the speech frames of a word from a backup dictionary. Thus, as noted in the prior response, Roberts' speech recognition system merely determines that a corrected word is not in the vocabulary and then isolates the speech frames for the new word from the speech frames for the utterance in order to build a speech model for the new word. Accordingly, the rejection should be withdrawn.

Applicant : Jonathan H. Young et al.
Serial No. : 09/390,370
Filed : September 7, 1999
Page : 8 of 8

Attorney's Docket No.: 13865-023001

Claim 36

Claim 36, which depends from claim 35, has been rejected as being unpatentable over Sherwood in view of Roberts and Baker. Applicant requests reconsideration and withdrawal of this rejection because Baker does not remedy the failure of Sherwood and Roberts to describe or suggest the subject matter of claim 35. In particular, like Sherwood and Roberts, Baker in no way describes or suggests generating the acoustic model for a word fragment that may be used to convert a word of an active vocabulary to a word of a backup dictionary by using a portion of an acoustic model of the word of the backup dictionary that is not included in an acoustic model of the word of the active vocabulary.

In the event that the Examiner feels that a telephonic or personal interview would be productive in advancing prosecution of this application, the Examiner is encouraged to telephone the undersigned at 202-626-6416.

No fees are believed to be due. Please apply any charges or credits to deposit account 06-1050.

Respectfully submitted,

Date: 4/29/04



John F. Hayden
Reg. No. 37,640

Fish & Richardson P.C.
1425 K Street, N.W.
11th Floor
Washington, DC 20005-3500
Telephone: (202) 783-5070
Facsimile: (202) 783-2331